

OCR 文字認識

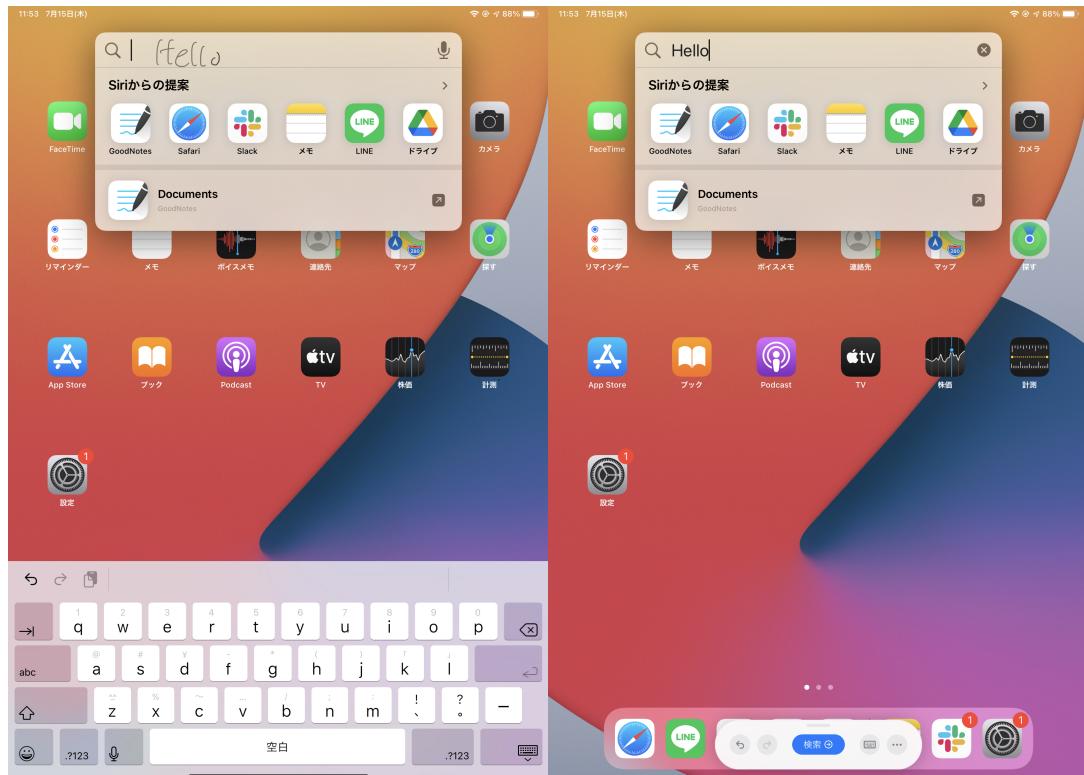
03-210245 峰岸剛基

概要

人間が書いた文字特に日本語を機械に認識させたい。

動機

iPad OS 14で導入されたスクリブルという手書きの文字をかなりの精度で認識する機能がある。下のように検索窓にApple pencil で書いた文字を検索できる。



しかしこまだ日本語対応していなく不便。だから自分で実装。

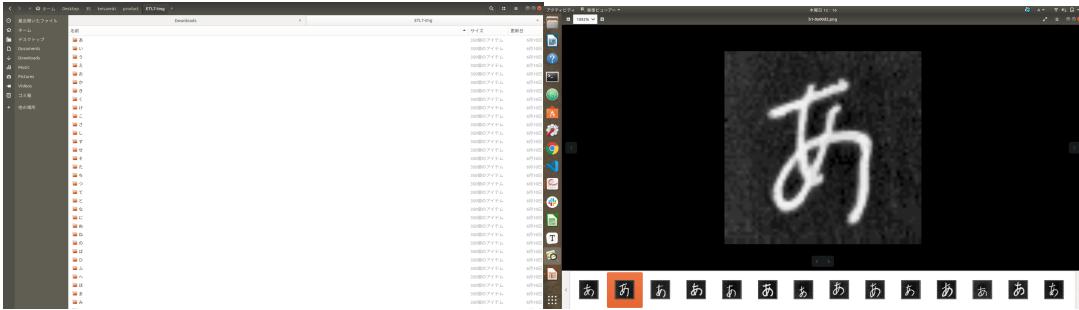
調査

調べてわかったこととして機械学習(Deep learning)を用いる必要があるということ。それには学習用の大量のデータが必要であるということ。

データ

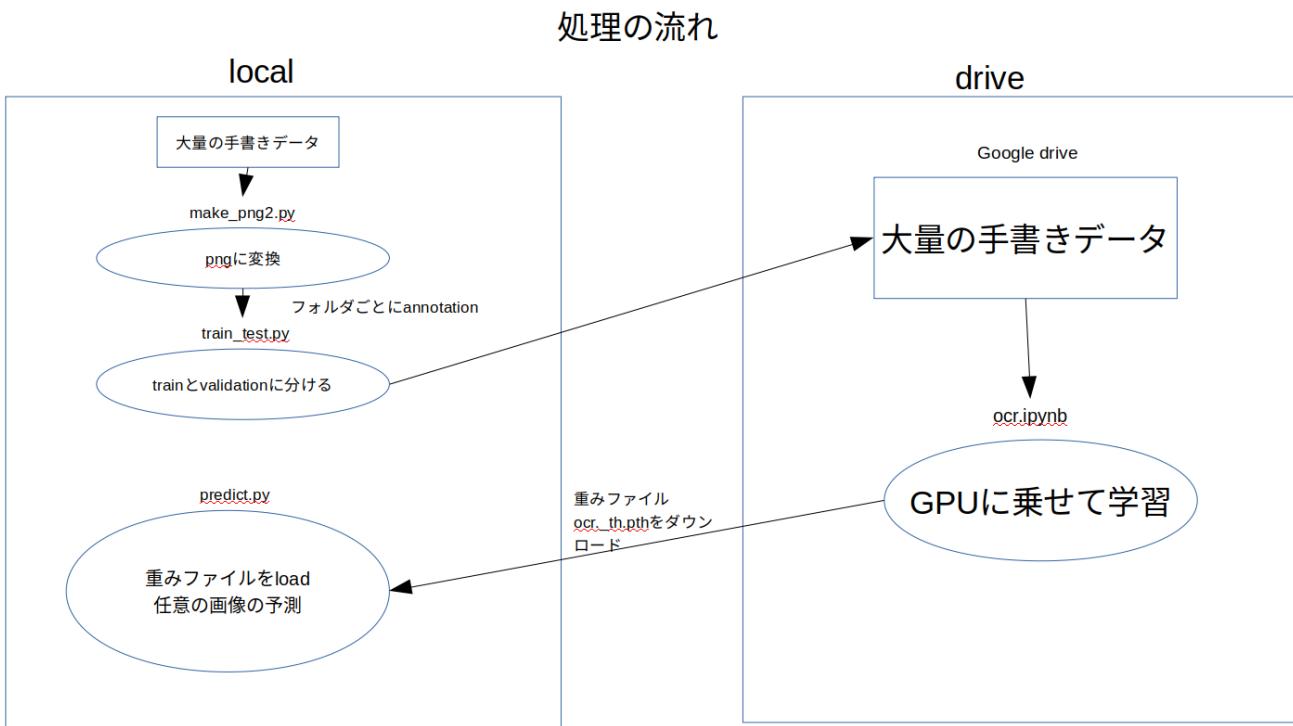
ETLデータベース

ひらがなと濁点と半濁点の合計48文字を認識する。漢字は多すぎて大変。謎の拡張子だったのでpngに変換。64*64pixelとか。



文字認識

学習にGPUを使わないと時間がかかるので無料でGPUが使えるGoogle colabで行った。学習後の重みファイルのみロードして実際に自分で書いた文字を予測する。



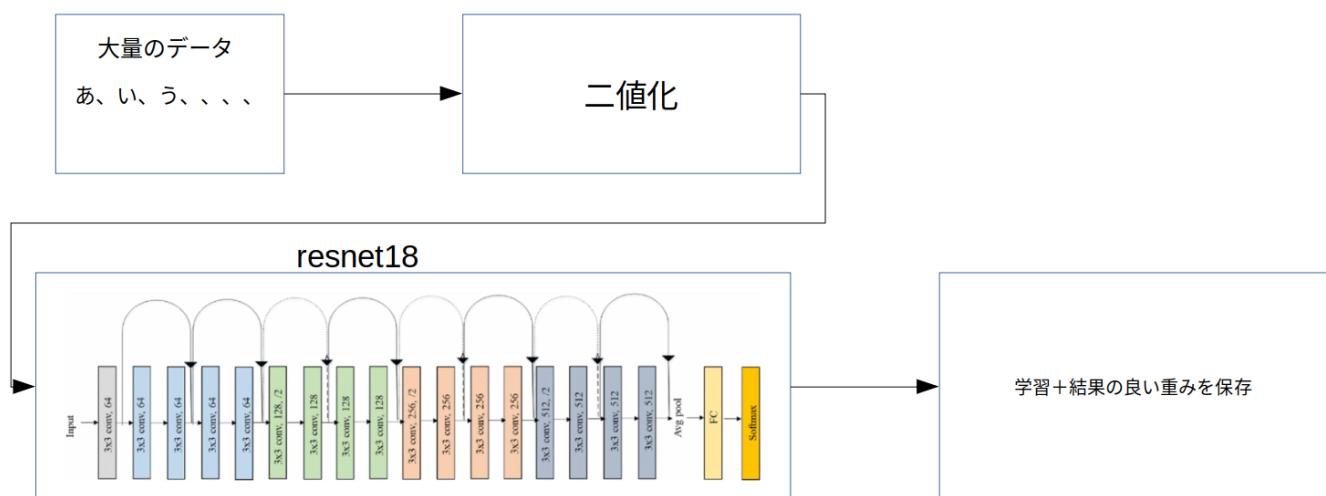
画像認識として精度のたかいResnet18を使用。

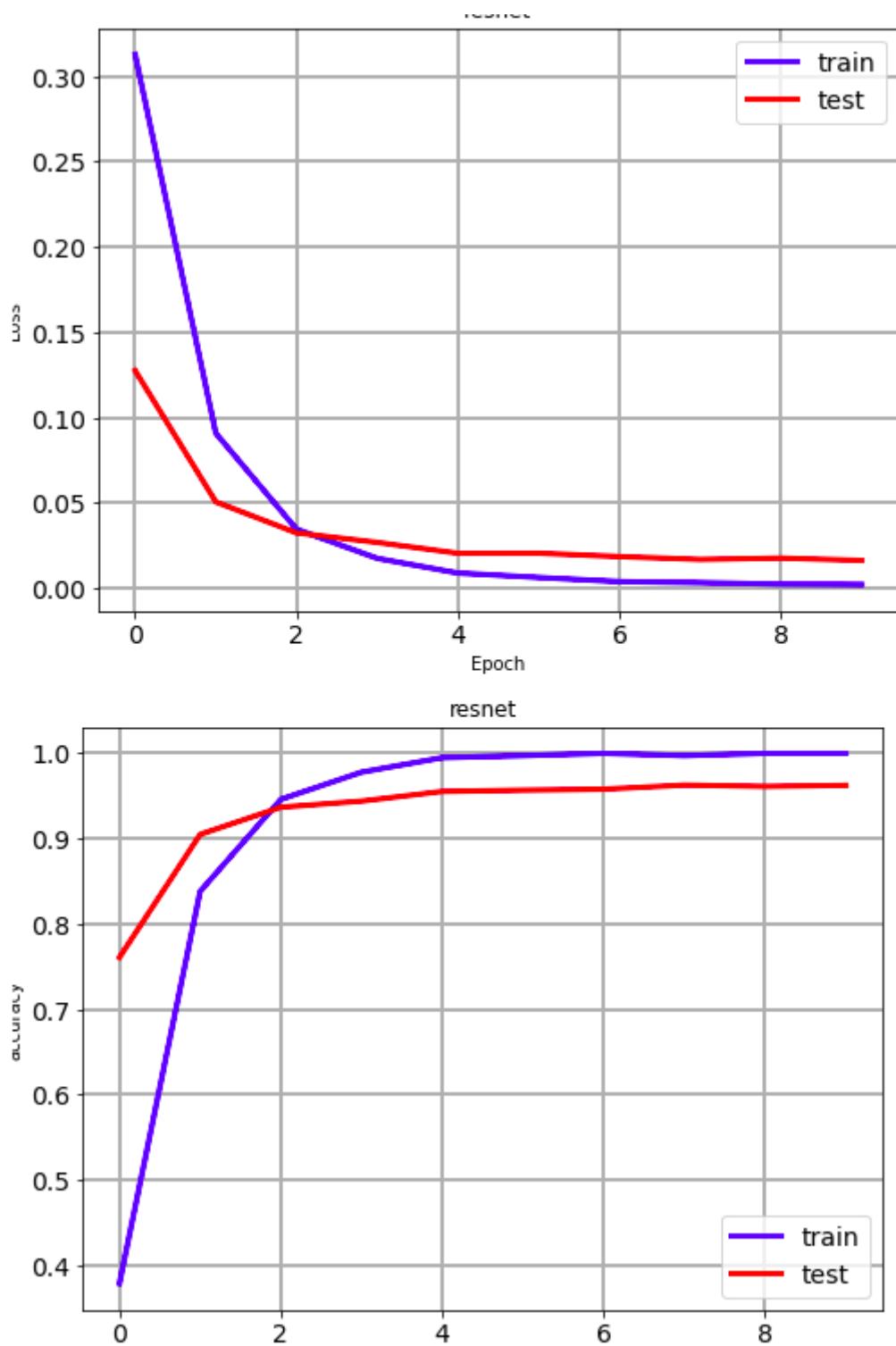
18は層の数を示す。

ディープラーニングでは基本的に層の数が多いほど複雑なものを表せるが今回二値化しているため18くらいでいいかと。あとcolabのGPUも多分12GBくらいしか使えないで、軽いモデルを選択した。

工夫した点は二値化したことと、Earlystopping で過学習を防止した。

モデルの概要





学習結果が以下

validationのaccuracyが95%くらいでした。

結果

データ iPadで書いた文字をスクショして切り取ったもの。

さ

出力

```
python predict.py 画像のpath
```

で画像のひらがながなにか予測

```
mech-user@KI00255: ~/Desktop/3S/keisannki/product
ファイル(F) 編集(E) 表示(V) 検索(S) 端末(T) ヘルプ(H)
$product mech-user@KI00255:~/Desktop/3S/keisannki/product$ python predict.py /home/mech-user/Desktop/3S/keisannki/product/handwritten/さ_1_ipad.jpg
さ
$product mech-user@KI00255:~/Desktop/3S/keisannki/product$
```

「さ」が認識できた！

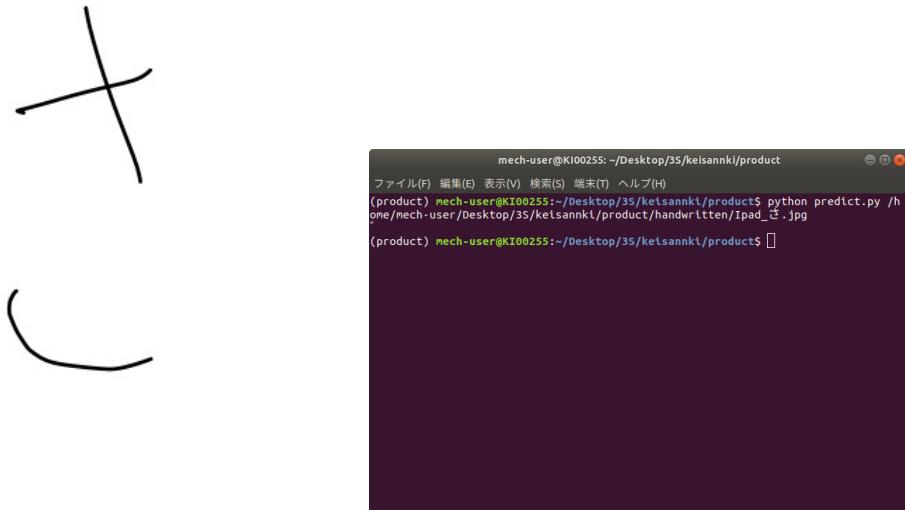
ちなみに他の文字や文字を太くしてもいけそう。

あ

```
mech-user@KI00255: ~/Desktop/3S/keisannki/product
ファイル(F) 編集(E) 表示(V) 検索(S) 端末(T) ヘルプ(H)
$product mech-user@KI00255:~/Desktop/3S/keisannki/product$ python predict.py /home/mech-user/Desktop/3S/keisannki/product/handwritten/あ_1_ipad.jpg
あ
$product mech-user@KI00255:~/Desktop/3S/keisannki/product$
```

反省

- 筆の太さがGood note の一番細いやつだとうまくできない。



- 精度が低い

以下は「あ、い、う、え、お」をそれぞれ予測させた結果である。

「え」が「ん」に認識されている。確かに似ているけど区別してほしかった。

```
mech-user@KI00255: ~/Desktop/3S/keisannki/product
ファイル(F) 編集(E) 表示(V) 検索(S) 端末(T) ヘルプ(H)
$product mech-user@KI00255:~/Desktop/3S/keisannki/product$ python predict.py /home/mech-user/Desktop/3S/keisannki/product/handwritten/あ.jpg
あ
$product mech-user@KI00255:~/Desktop/3S/keisannki/product$ python predict.py /home/mech-user/Desktop/3S/keisannki/product/handwritten/い.jpg
い
$product mech-user@KI00255:~/Desktop/3S/keisannki/product$ python predict.py /home/mech-user/Desktop/3S/keisannki/product/handwritten/う.jpg
う
$product mech-user@KI00255:~/Desktop/3S/keisannki/product$ python predict.py /home/mech-user/Desktop/3S/keisannki/product/handwritten/え.jpg
え
$product mech-user@KI00255:~/Desktop/3S/keisannki/product$ python predict.py /home/mech-user/Desktop/3S/keisannki/product/handwritten/お.jpg
お
$product mech-user@KI00255:~/Desktop/3S/keisannki/product$ 
```

- 複数の文字をは無理

授業でもやったがopencvで微分とかして輪郭求めようとしてもしたの画像のようになってしまふ。文字の位置を認識するモデルも必要みたい。



- IPad OS 15でスクリブルが日本語対応してしまうらしい
- もっとみんなみたいにかっこいいわかりやすいUIを作って「すごい」って思われたかった。